

# **Introduction à l'acoustique de la parole**

*Yves Laprie*

## Présentation

- **Énoncé du problème acoustique**
  - Production des sons de la parole à partir d'une fonction d'aire (voyelles et dans une moindre mesure consonnes),
  - Inversion impossible
- **Modèles articulatoires**
  - définir une fonction d'aire à partir d'un nombre restreint de paramètres,
  - Synthèse articulatoire,
  - Codage articulatoire,
  - Contrôle dynamique impossible,
- **Phonologie articulatoire et contrôle musculaire**
  - Activation des articulateurs,
  - Modélisation de la coarticulation
- **Inversion acoustique articulatoire**

# 1 Acoustique

Qu'apporte la résolution des équations de l'acoustique ?

- Évaluer la conséquence acoustique des variations des paramètres physiques du conduit vocal (forme, pertes, rigidité des parois...)
- Étude du couplage sous-glottique
- Étude théorique de la production des sons : voyelles, occlusives, fricatives
- Étude de la source pour les consonnes
- Synthèse de parole

## Équation de propagation d'une onde plane

problème : lier la pression, la vitesse et la densité de l'air

Pour obtenir quoi ?

- la connaissance des fonctions  $p(x, t)$ ,  $v(x, t)$  et  $\rho(x, t)$ ,
- le rapport débit d'air à la sortie du conduit vocal sur le débit à la glotte, c'est-à-dire la fonction de transfert ou le spectre.

Avec quelles données :

- la fonction d'aire  $A(x, t)$ , plus souvent  $A(x)$ ,
- le débit d'air à la glotte, voire la pression des poumons et la variation temporelle de l'aire entre les cordes vocales.

Il faut donc 3 équations liant  $v$ ,  $p$  et  $\rho$ .

## Équation d'état

On considère que :

- la propagation du son est adiabatique,
- le gaz est un gaz parfait.

$p$  et  $\rho$  sont liées par :

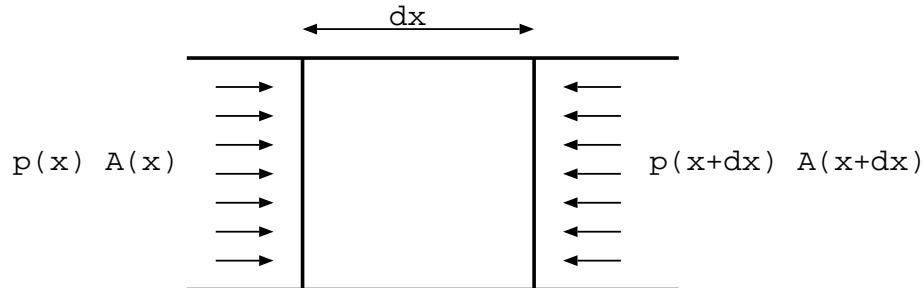
$$\rho = \rho_0 + \frac{1}{c^2} p$$

et donc leurs dérivées sont liées par :

$$\boxed{\frac{\partial \rho}{\partial t} = 1/c^2 \frac{\partial p}{\partial t}}$$

## Équation d'Euler

somme des forces appliquées = masse  $\times$  accélération



bilan des forces :

$$-(p(x + dx, t)A(x + dx) - p(x, t)A(x)) = -A(x) \frac{\partial p}{\partial x} dx$$

accélération :

$$m\gamma = dx \rho(x, t) A(x) \frac{dv}{dt}$$

or  $dv = \frac{\partial v}{\partial t} dt + \frac{\partial v}{\partial x} dx$  soit  $\frac{dv}{dt} = \frac{\partial v}{\partial t} + v \frac{\partial v}{\partial x}$  avec  $v \frac{\partial v}{\partial x}$  petit. Donc

$$m\gamma = \rho(x, t) A(x) dx \frac{\partial v}{\partial t} \approx \rho_0 A(x) dx \frac{\partial v}{\partial t}$$

L'équation d'Euler est donc :

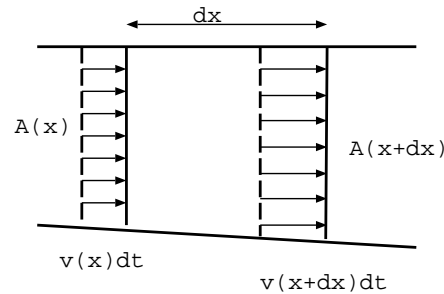
$$\boxed{-\frac{\partial p}{\partial x} = \rho_0 \frac{\partial v}{\partial t}}$$

approximations :

- grandeurs acoustiques  $p_a = p - p_{atmosphérique}$ ,  $v_a = v - v_{gaz}$ ,
- pression acoustique : de l'ordre du Pa ( $\ll$  pression atmosphérique =  $10^5$  Pa),
- vitesse  $10^{-7} m s^{-1}$  au seuil d'audition,
- les dérivées de fonctions sinusoïdales sont du même ordre de grandeur que les fonctions (vitesse, pression),
- aire : de l'ordre du  $cm^2$  et varie peu en fonction de x.

## Équation de continuité

flux de matière = taux d'accroissement de matière d'une tranche



**excès de matière pour une tranche pendant dt :**

$$\begin{aligned}
 & (A(x)\rho(x, t)v(x, t) - A(x + dx, t)\rho(x + dx, t)v(x + dx, t))dt \\
 &= -\left(\frac{\partial A}{\partial x}\rho(x, t)v(x, t) + \frac{\partial \rho}{\partial x}A(x)v(x, t) + \frac{\partial v}{\partial x}A(x)\rho(x, t)\right)dxdt \\
 &\text{or } \frac{\partial \rho}{\partial x}A(x)v(x, t) \text{ est négligeable car } v(x) \text{ et } \frac{\partial \rho}{\partial x} \text{ sont petits. Le flux est donc :} \\
 & -\rho(x, t)\left(\frac{\partial A}{\partial x}v(x, t) + \frac{\partial v}{\partial x}A(x)\right)dxdt \\
 &\text{ou encore négligeant les termes d'ordre 2 dus à } \rho(x, t) = \rho_0 + d\rho : \\
 & -\rho_0\left(\frac{\partial A}{\partial x}v(x, t) + \frac{\partial v}{\partial x}A(x)\right)dxdt
 \end{aligned}$$

**taux d'accroissement de matière :**

$$\frac{\partial A(x)\rho(x, t)}{\partial t} dt dx = A(x) \frac{\partial \rho(x, t)}{\partial t} dt dx$$

En remplaçant  $\frac{\partial \rho}{\partial t}$  grâce à l'équation d'état :



$$= \frac{A(x)}{c^2} \frac{\partial p(x,t)}{\partial t} dt dx$$

D'où l'équation de continuité :

$$-\rho_0 \left( \frac{\partial A}{\partial x} v(x,t) + \frac{\partial v}{\partial x} A(x) \right) = \frac{A(x)}{c^2} \frac{\partial p(x,t)}{\partial t}$$

## Équation de Webster

Par dérivation par rapport à  $t$  de l'équation de continuité on obtient :

$$-\rho_0 \left( A \frac{\partial^2 v}{\partial x \partial t} + \frac{\partial v}{\partial t} \frac{\partial A}{\partial x} \right) = \frac{A}{c^2} \frac{\partial^2 p}{\partial t^2}$$

En remplaçant  $\frac{\partial v}{\partial t}$  grâce à l'équation d'Euler on obtient :

$$A \frac{\partial^2 p}{\partial x^2} + \frac{\partial p}{\partial x} \frac{\partial A}{\partial x} = \frac{A}{c^2} \frac{\partial^2 p}{\partial t^2}$$

ou encore la célèbre équation de Webster :

$$\boxed{\frac{1}{A(x)} \frac{\partial}{\partial x} \left[ A(x) \frac{\partial p}{\partial x} \right] = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}}$$

- pas de résolution analytique sauf fonction d'aire très spécifique (tuyau conique, parabolique, hyperbolique...).
- pas de perte.

## Équations de l'acoustique

équation d'Euler :

$$\boxed{-\frac{\partial p}{\partial x} = \rho_0 \frac{\partial v}{\partial t}}$$

équation de continuité :

$$\boxed{-\rho_0 \left( \frac{\partial A}{\partial x} v(x, t) + \frac{\partial v}{\partial x} A(x) \right) = \frac{A(x)}{c^2} \frac{\partial p(x, t)}{\partial t}}$$

équation de Webster :

$$\boxed{\frac{1}{A(x)} \frac{\partial}{\partial x} \left[ A(x) \frac{\partial p}{\partial x} \right] = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}}$$

et sans variation d'aire :

$$\begin{cases} \frac{\partial^2 v}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 v}{\partial t^2} \\ \frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2} \end{cases}$$

## Prise en compte de la spécificité du conduit vocal

**Vibration de la paroi**  $m\ddot{y} + b\dot{y} + k(y - y_0) = p(x, t)$  avec

$$A(x, t) = A_0(x, t) + y(x, t)S_0(x, t)$$

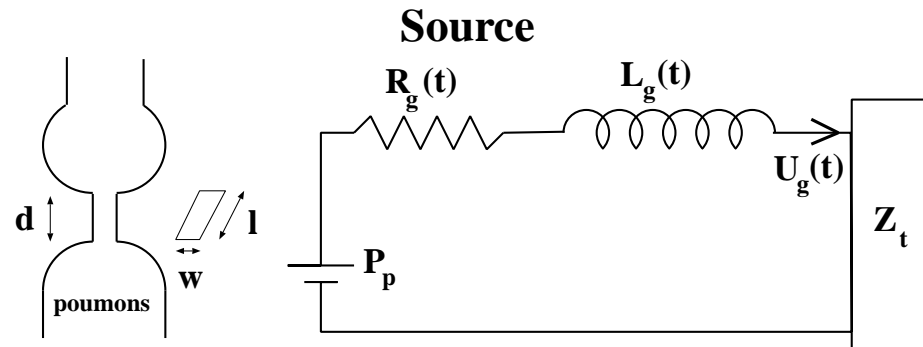
**Conduit dynamique**  $A = A(x, t)$  donc  $\frac{\partial A}{\partial t} \neq 0$

**Conditions aux limites** •  $p(x = 0, t) = P_{subglottique}$

- couplage nasal : conservation du débit, continuité de la pression.

**Radiation aux lèvres**

**Pertes** adjonction d'un terme  $Rv/A_0$  dans l'équation d'Euler.



$U_g(t)$  est créé par la pression des poumons modulée par la variation de l'aire à la glotte  $A_g(t)$ .

Deux phases : une ouverte, une fermée.

$$\begin{cases} R_g(t) = \frac{12\mu d}{lw^3} + 0.875 \frac{\rho U_g(t)}{2(lw)^2} \\ L_g(t) = \rho \frac{d}{A(t)} \end{cases}$$

Équation différentielle électrique :

$$U_g(t)R_g(t) + \frac{d}{dt}[L_g(t)U_g(t)] = P_p$$

Résolution en simplifiant ou en utilisant la méthode de Runge-Kutta. Autre méthode : 2 masses vibrantes (Ishizaka).

## Résolution des équations

### Résolution temporelle

- discrétisation dans le temps et dans l'espace,
- équations aux différences,

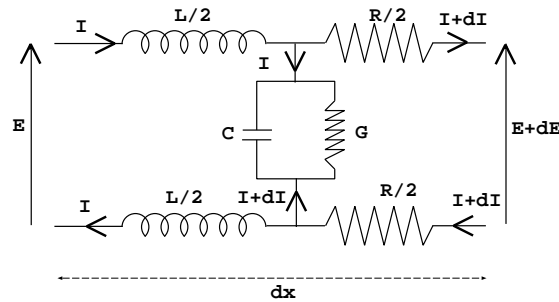
Avantage : prise en compte de l'interaction source conduit

### Résolution fréquentielle

- avec l'analogie électrique,
- découpage du conduit en tubes élémentaires.

Avantage : rapide, modélisation séparée du conduit et de la source.

## Analogie électrique



- quadripôle électrique,
- $R, L, C$  et  $G$  résistance, impédance, capacitance et conductance par unité de longueur,
- on cherche  $I$  et  $E$  sous la forme  $f(x)e^{j\omega t}$ .

On pose  $z = R + jL\omega$  (en  $\Omega m^{-1}$ ) et  $y = C + j\omega C$  (en  $\Omega^{-1} m^{-1}$ ).

$$\begin{cases} E = \frac{z}{2} dx I - \frac{dI}{y dx} \\ E + dE = -(I + dI) \frac{z}{2} dx - \frac{dI}{y dx} \end{cases} \quad \begin{cases} E dx = \frac{z}{2} dx^2 I - \frac{dI}{y} \\ dE = -z I dx - \frac{z}{2} dx dI \end{cases}$$

On obtient donc :

$$\begin{cases} dI = -E y dx \\ dE = -I z dx \end{cases} \quad \text{soit} \quad \begin{cases} \frac{d^2 I}{dx^2} = y z I \\ \frac{d^2 E}{dx^2} = y z E \end{cases}$$

En revenant à l'acoustique et en supposant  $p(x, t)$  de la forme  $p(x)e^{j\omega t}$  l'équation  $\frac{\partial^2 p}{\partial x^2} = \frac{1}{c^2} \frac{\partial^2 p}{\partial t^2}$  s'écrit :

$$\frac{d^2 p(x)}{dx^2} e^{j\omega t} = -\frac{\omega^2}{c^2} p(x) e^{j\omega t}$$

Ce qui montre l'analogie entre les problèmes.



## Résolution fréquentielle

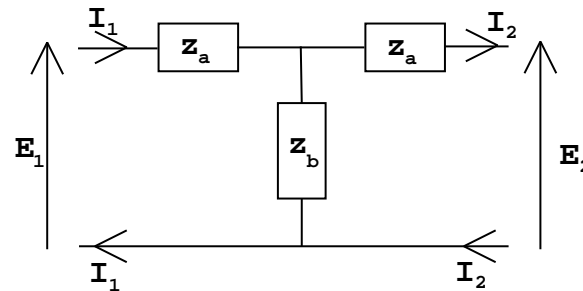
On cherche  $E(x)$  et  $I(x)$  sous la forme :

$$\begin{cases} E = A_1 e^{\gamma x} + B_1 e^{-\gamma x} \\ I = A_2 e^{\gamma x} + B_2 e^{-\gamma x} \end{cases}$$

avec  $\gamma = \sqrt{yz}$ .

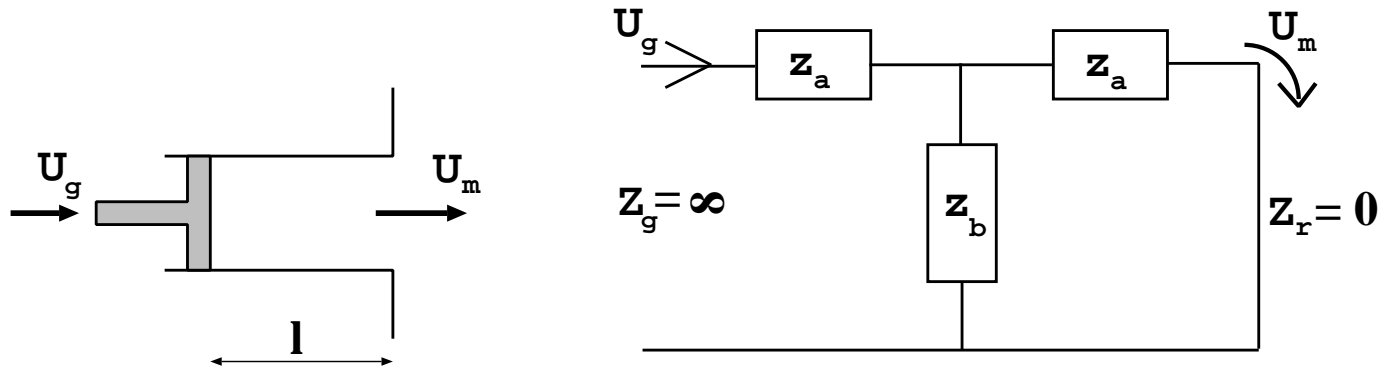
Détermination des constantes par l'expression des conditions aux limites sur  $E(x = 0)$ ,  $I(x = 0)$ ,  $\frac{dE}{dx}(x = 0)$  et  $\frac{dI}{dx}(x = 0)$ .

Et Analogie avec un quadripôle en T :



avec  $z_a = Z_0 \operatorname{th} \frac{\gamma l}{2}$  et  $z_b = Z_0 \frac{1}{\operatorname{sh} \gamma l}$  et  $Z_0 = \rho c / A$

### Exemple : le conduit uniforme



Sur la boucle droite on a :

$$z_a U_m = (U_g - U_m) z_b \text{ soit (en se rappelant que } th \frac{\gamma l}{2} = \frac{ch\gamma l - 1}{sh\gamma l} \text{):}$$

$$\frac{U_m}{U_g} = \frac{z_b}{z_a + z_b} = \frac{1}{ch\gamma l}$$

Sans perte, c'est-à-dire avec  $\gamma = j\omega\sqrt{LC} = j\omega/c$  le dénominateur s'annule pour :

$$\omega = (2n + 1) \frac{\pi c}{2l}$$

ou :  $f_n = (2n + 1) \frac{c}{4l}$  avec  $c = 350 \text{ms}^{-1}$  et  $l = 17 \text{cm}$  cela donne les formants :

$$f_1 = 514 \text{Hz}, f_2 = 1542 \text{Hz}, f_3 = 2573 \text{Hz} \dots$$

## Conduits à plus d'un tube

- 1 tube → /œ/
- 2 tubes → /a, i/
- 4 tubes avec 3 paramètres → à peu près tous les sons permet de calculer des nomogrammes.
- tubes en parallèle pour étudier le couplage nasal.

## Problème inverse

- une infinité de fonctions d'aire peuvent produire le même spectre !
- 17 tranches pour 3 formants !
- adopter un modèle avec moins de tubes permet de résoudre mais la solution n'est plus réaliste.
- problème de la source
- techniques de “tri” dans des ensembles de fonctions d'aires (Atal 78).

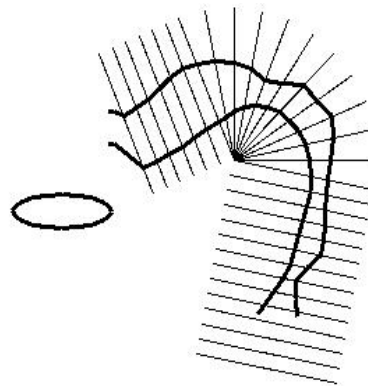
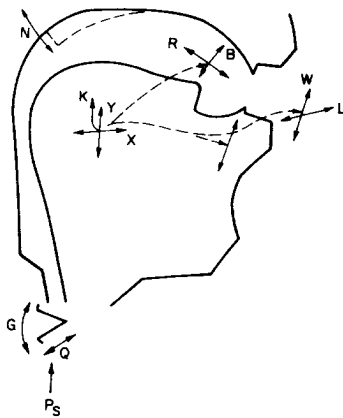
Comment réduire le nombre de paramètres en conservant des fonctions d'aire réalistes ?

## 2 Modèles articulatoires

But : obtenir une fonction d'aire à partir d'un ensemble restreint de paramètres dans le cadre de la synthèse articulatoire.

**modèles géométriques** Coker, Mermelstein... Modèles simples mais les paramètres ne sont pas forcément réalistes.

**modèles anthropomorphes** Maeda, Gabioud... Modèles obtenus à partir de l'analyse d'images cinéradiographiques.



**paramètres :**  
position du maxillaire  
position du dos de la langue  
forme du dos de la langue  
position de l'apex  
hauteur des lèvres  
protrusion des lèvres  
hauteur du larynx

## Construction du modèle de Maeda

Utilisation de coordonnées semi-polaires.

- **Analyse en composantes principales** → 3 composantes expliquent 98% de la variance.
- **Analyse générale en composantes linéaires** On maîtrise le choix des facteurs. Soit  $Z$  le vecteur des points définissant la langue,  $f_i$  l'influence du  $i$ ème facteur,  $a_p$  le vecteur des poids des différents facteurs sur la langue.  $z = \sum_i \sum_p a_{pi} f_i$   
Pour chaque facteur on calcule sa corrélation avec les formes mesurées, ce qui donne le vecteur de poids pour ce facteur.  
On soustrait la variance expliquée par le facteur et on recommence avec le facteur suivant.

## Codage articulatoire

But : compresser le signal de parole dans le but de transmettre à moindre coût.

Algorithme :

- création d'un dictionnaire de formes articulatoires indexé acoustiquement.
- récupérer à intervalles régulier les formes du conduit vocal (choix du type de parole, amélioration des paramètres. . .).
- naviguer en minimisant une contrainte de variation de forme sur le segment de parole à coder.

La contrainte d'évolution permet de réduire l'indéterminisme.

Dans la pratique : un Cray Ymp et un locuteur.

## Décodage acoustico-phonétique articulatoire

Difficulté : choisir les cibles

acoustique – articulatoire – constriction – autre ?

→ définition de patrons articulatoires et comparaison acoustique.

→ mélange acoustique articulatoire :

**simple mélange** vecteur acoustique articulatoire dans un HMM,

### apprentissage de l'inversion

- faire parler un locuteur devant un micro et avec un articulographe (EMA) → entraîner un réseau neuromimétique à récupérer l'articulatoire à partir de l'acoustique puis un HMM avec acoustique plus articulatoire.
- après apprentissage utiliser le HMM avec l'articulatoire fourni par le réseau neuromimétique.

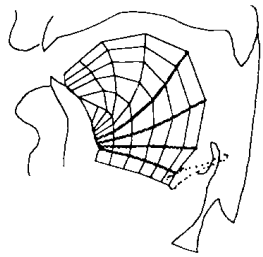


## Contrôle musculaire du conduit vocal

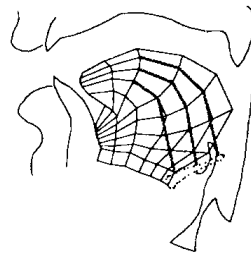
But : déterminer les paramètres articulatoires à partir de l'activation des muscles du conduit vocal

- modélisation par éléments finis des articulateurs,
- équation différentielle du second ordre pour contrôler les muscles :

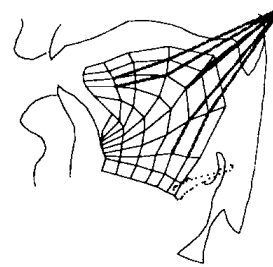
$$m\ddot{x} + b\dot{x} + k(x - x_0) = 0$$



**Génioglosse  
postérieur**



**Hyoglosse**



**Styloglosse**

### 3 Phonologie articulatoire

But : prédire les gestes articulatoires à partir de la connaissance des phonèmes

- geste  $\Leftrightarrow$  formation ou relâchement d'une constriction,
- but du geste déterminé en termes de variables dimensionnelles du conduit vocal,
- une variable est associée à un ensemble d'articulateurs  
ouverture des lèvres : maxillaire, lèvre inf, lèvre sup
- mouvement d'une variable définie par un système dynamique du second ordre :  
$$m\ddot{x} + b\dot{x} + k(x - x_0) = 0$$

## Scores articulatoires

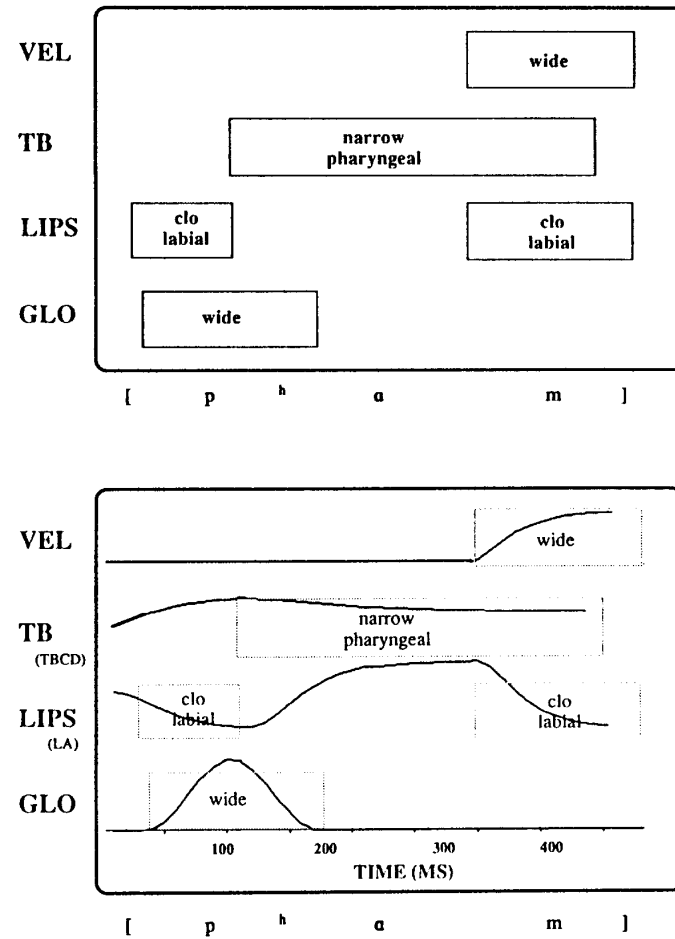


Figure 3

Gestural score for *palm* [pɑm] using box notation. (a) Activation intervals only; (b) model generated tract variable motions added.

## Organisation des gestes articulatoires

- il peut y avoir recouvrement des gestes,
- il faut spécifier leur organisation temporelle,
- or activation un articulateur est attiré par une position neutre,
- deux gestes peuvent :
  - porter sur des variables articulatoires différentes
    - \* les deux gestes ont lieu et sont perceptibles,
    - \* un geste peut cacher acoustiquement l'autre.
  - se mélanger (somme des variables dans le temps).

## Topologie d'un HMM et phonologie articulatoire

Idée : Dans un HMM la variabilité passe dans la variance  
réduire la variance grâce à l'articulatoire

- faire appel à la phonologie articulatoire pour construire la topologie des modèles,
- chaque phonème est décrit en termes de variables articulatoires (Lèvres, Apex, Corps de la langue, Velum, Larynx),
- Prédiction des mouvements observés en fonction des description articulatoires.

### Construction des états

	t	ε	n	k	æ	t	s
<b>lèvres</b>	0	0	0	0	0	0	0
<b>apex</b>	1	0	1	0	0	1	3
<b>langue</b>	0	9	0	1	10	0	0
<b>voile</b>	1	1	2	1	1	1	1
<b>larynx</b>	2	1	1	2	1	2	2

- algorithme de construction des automates utilisant des contraintes phonétiques,
- apprentissage et reconnaissance classique.

## 4 Inversion acoustique articulatoire

Pourquoi inverser la parole ?

- Codage de la parole
  - petit nombre de paramètres (source + forme du conduit vocal)
  - fidèle à la production de parole
- Reconnaissance automatique de la parole
  - compléter voire remplacer le signal acoustique par un signal articulatoire (analogie avec la reconnaissance audio-visuelle)
  - meilleure prise en compte de la coarticulation (articulateur critiques)
  - définition de cibles articulatoires (consonnes) ou de cibles acoustiques (voyelles)
- Apprentissage de la parole (langues étrangères)
- Evaluation des pathologies de l'articulation

## Difficultés de l'inversion

Il existe une infinité de formes articulatoires qui donnent le même spectre de parole (cf. Atal 78) !

Les deux solutions extrêmes :

1. **Conduit vocal peu contraint + techniques de tabulation complétées par des contraintes (Atal et al. 78, Scroeter et al. 93)**
  - Avantage : phénomènes compensatoires conservés
  - Désavantage : nécessité d'un dictionnaires de formes
2. **Conduit vocal très contraint + solution presque analytique (Schoentgen et al. 95)**
  - Avantage : précision des solutions ( $< 0.01$  Hz par formant)
  - Désavantage : peu écologique



### **Solution généralement adoptée pour l'inversion**

- Conduit vocal produit par un modèle articulatoire (le conduit est donc relativement réaliste)
- Utilisation de contraintes portant sur l'évolution du conduit vocal (pas trop rapide, pas trop extrême, pas trop chaotique)

## Méthodes par tabulation

### 1. Construction d'un dictionnaire de formes articulatoires

Une forme articulatoire donnée par  
ses paramètres articulatoires

$\Leftrightarrow$

Les 3 premiers for-  
mants

**Problème :**

comment trouver les bonnes formes articulatoires ?

### 2. Construction des trajectoires articulatoires

À chaque instant on recherche les formes qui correspondent aux formants F1 F2 F3

**Problème :** comment assurer que les trajectoires sont réalisables par un locuteur humain ?

## Notre approche

- Objectifs :**
- utiliser le comportement acoustique du modèle articulatoire durant l'inversion conserver les phénomènes compensatoires
  - prendre en compte simultanément la distance acoustique et l'évolution des paramètres articulatoires.

- Choix :**
- modèle articulatoire de Maeda
  - travailler à partir des formants

- Notations**
- paramètres articulatoires  $\alpha(t) = (\alpha_1(t) \dots \alpha_i(t) \dots \alpha_7(t)) \quad t \in [t_i, t_f]$
  - formants générés par la simulation acoustique  $F_j(\alpha(t)) \quad 1 \leq j \leq 3$
  - formants extraits de la parole ou mesurés  $f_j(t) \quad 1 \leq j \leq 3$

## Construction de trajectoires articulatoires

1. Récupérer les forme qui correspondent à un spectre
  - à partir du spectre directement
    - distance acoustique et perceptive
  - à partir de paramètres extraits du spectre
    - pics ou formants
2. Trouver des trajectoires articulatoires lisses
  - programmation dynamique
    - minimiser l'énergie cinétique et potentielle  $\frac{1}{2} \sum_i m_i \alpha_i'^2 + \frac{1}{2} \sum_i k_i \alpha_i^2$
  - méthodes neuromimétiques
3. Optimisation finale
  - améliorer la proximité acoustique entre les résultats et le signal de départ
    - jouer sur les paramètres de source
    - danger : perdre la régularité des trajectoires

## Détermination des solutions initiales

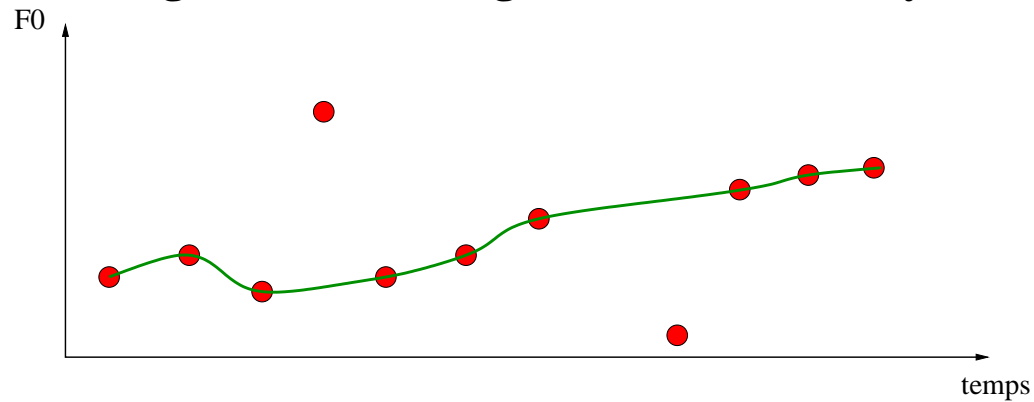
Il faut des solutions initiales assez proches de solutions réalistes.

=> Améliorer une méthode d'inversion par tabulation On peut approcher acoustiquement la parole initiale très précisément mais au prix de trajectoires très chaotiques.

1. échantillonnage articulaire insuffisant
2. proximité d'une non linéarité acoustique
3. Modèle articulaire incapable de produire une forme de conduit identique à celle réalisée par le locuteur ? effet compensatoire injustifié

=> construction de solutions initiales éventuellement partielles grâce à un algo de lissage proposé par Ney.

### Algorithme de lissage non linéaire de Ney



$C = (c(0), c(1) \dots c(i)c(N))$  points de départ Idée : choisir parmi les points de départ un sous-ensemble de points formant une courbe suffisamment lisse.

$$\bar{C} = [c(j(k))] = c(j(1)) \dots c(j(k)) \dots c(j(K))$$

Avec  $K \leq N$  et  $j$  fonction d'indice strictement croissante :  $0 \leq j(k) \leq N$ .

## Résolution

- Minimisation du critère global

$$D = \sum_{k=1}^{K-1} d(c(j(k)), c(j(k+1)))$$

$d$  est une distance entre 2 points

**Mais  $D$  est minimal avec un sous-ensemble vide !**

- Adjonction d'un bonus

$$D = \sum_{k=1}^{K-1} d(c(j(k)), c(j(k+1))) - \text{Bonus}(c(j(k)))$$

**B fort : beaucoup de points B faible : peu de points**

- Résolution par la programmation dynamique

## Construction du dictionnaire de formes articulatoires

Problème : comment échantillonner les paramètres articulatoires ?

7 paramètres

3 écarts type autour de la moyenne, pas de 0,1?

$\Leftrightarrow$

$30^7$  formes

Echantillonnage aléatoire

- choisir aléatoirement 100000 formes
- vérifier que les fonctions d'aire sont réalistes (constriction pas trop fortes)

Utilisation de formes de base

- formes articulatoires des voyelles
- échantillonnage autour des trajectoires entre deux formes de base quelconques



## Méthode variationnelle

Fonction de coût pour évaluer l'inversion :

$$\begin{aligned}
 I = & \underbrace{\int_{t_i}^{t_f} \sum_{j=1}^3 (f_j(t) - F_j(\alpha(t)))^2 dt}_{\text{Distanceacoustique}} \\
 & + \lambda \underbrace{\int_{t_i}^{t_f} \sum_{i=1}^7 m_i \alpha_i'^2(t) dt}_{\text{Energiecinétique}} + \beta \underbrace{\int_{t_i}^{t_f} \sum_{i=1}^7 k_i \alpha_i^2(t) dt}_{\text{Energiepotentielle}}
 \end{aligned} \tag{1}$$

Forme de I:  $I = \int_{t_i}^{t_f} \Phi(\alpha(t), \alpha'(t), t) dt$

Exprimer l'annulation de la dérivée de I par rapport aux fonctions

$$\left\{ \begin{array}{l} \frac{\partial \Phi}{\partial \alpha_1} - \frac{d}{dt} \frac{\partial \Phi}{\partial \alpha_1'} = 0 \\ \dots \\ \frac{\partial \Phi}{\partial \alpha_7} - \frac{d}{dt} \frac{\partial \Phi}{\partial \alpha_7'} = 0 \end{array} \right. \tag{2}$$

## Travail préalable : adaptation du modèle articulatoire de Maeda

- Modèle de Maeda obtenu à partir d'images d'une locutrice
- Modèle capable de produire les sons d'un seul locuteur

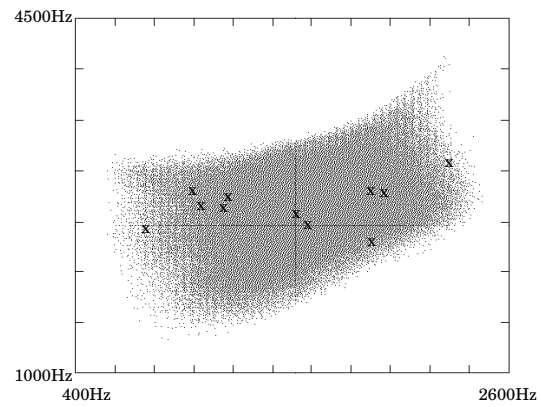
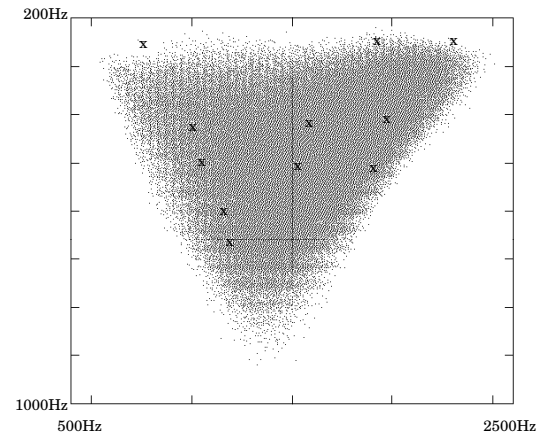
Adapter le modèle :

- taille de la bouche, taille du pharynx, paroi fixe
- **mais le comportement articulatoire reste le même.**

Une série d'images IRM pour 11 voyelles :

- le modèle de Maeda a été obtenu sur des images aux rayons X !
- l'IRM provoque Effet Lombard.
- Taille de la bouche x1.08
- Taille du pharynx x1.18
- Paroi obtenue par moyenne puis optimisation
- Erreurs commises sur les voyelles d'adaptation : 49 Hz pour F1, 125 Hz pour F2, 170 Hz pour F3

## Adaptation

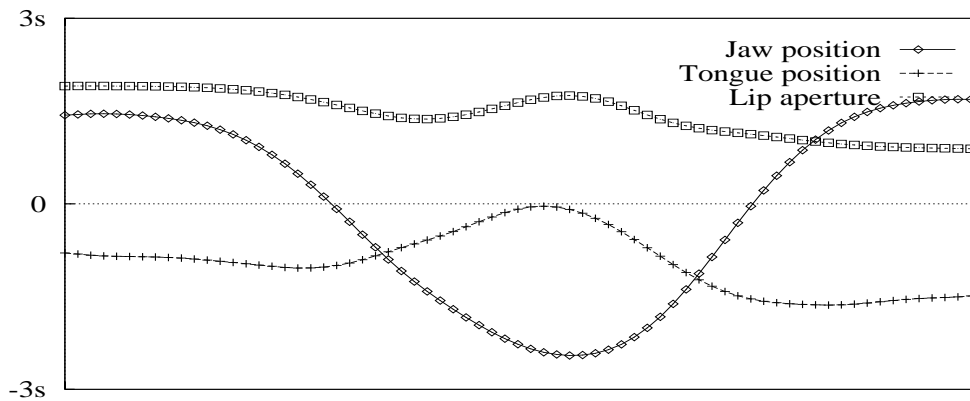
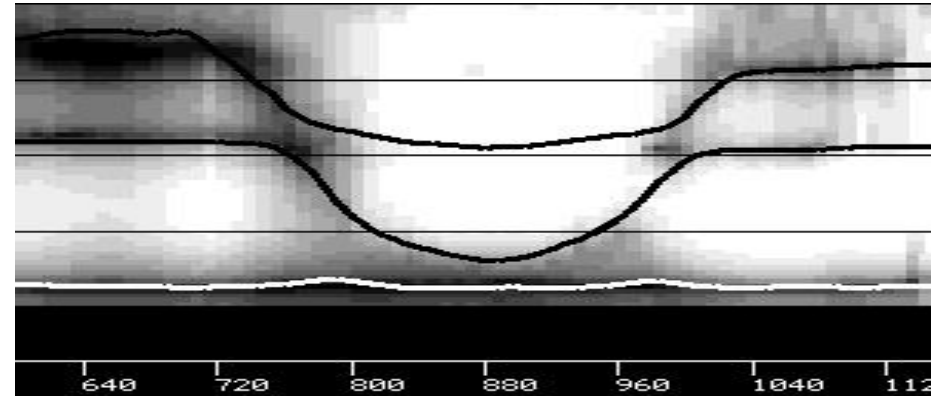
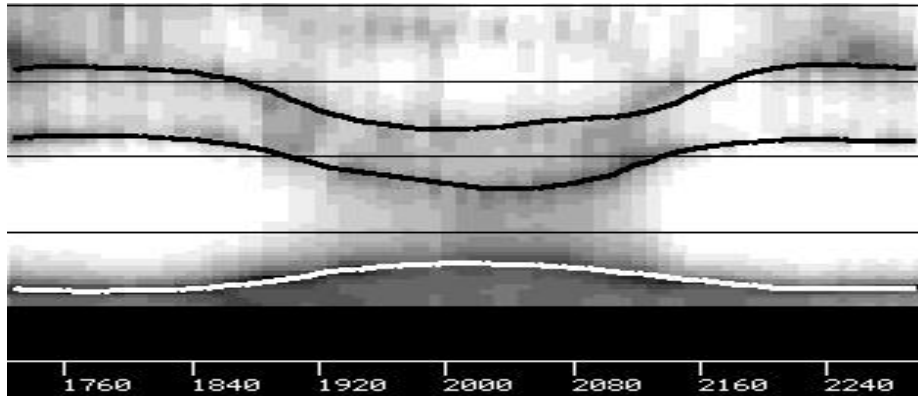


Modèle de Maeda recalé et espaces vocaliques générés (F1-F2 en haut, F2-F3 en bas). les croix représentent les voyelles du sujet

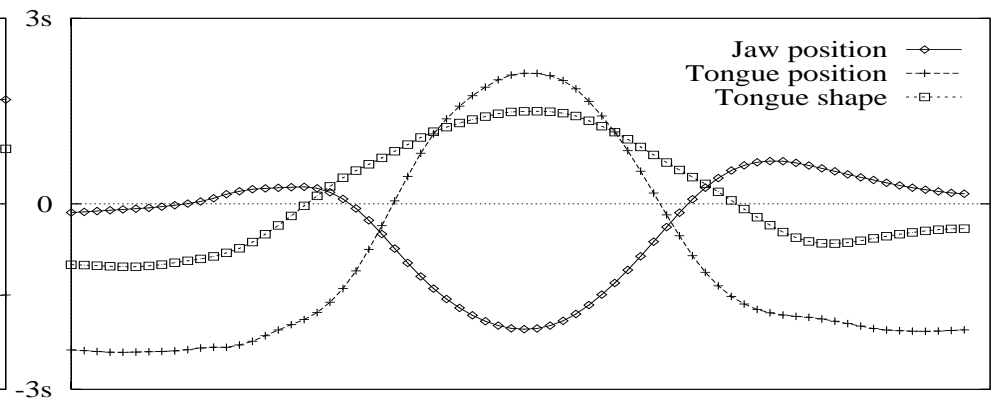
## Expériences

- Construction des dictionnaires de formes
  - échantillonnage aléatoire
  - formes de base
    - Première solution car hélas les voyelles ont été centralisées à cause du bruit de la machine IRM
  - 300000 formes filtrées par l'aire de leur constriction principale.
- Hypothèses simplificatrices :  
constantes  $m_i$  et  $k_i$  fixées à 1 (même inertie pour la mâchoire et les lèvres)
- Inversion
  - construction des solutions initiales
  - optimisation des trajectoires par l'approche variationnelle

## Expériences



Trajectoires formantiques générées à partir des paramètres articulatoires récupérés pour /ai/ (paramètres donnés par rapport à la position neutre).



Trajectoires formantiques générées à partir des paramètres articulatoires récupérés pour /iui/ (paramètres donnés par rapport à la position neutre).

## Améliorations possibles

- Revenir sur les simplifications abusives (mi, ki?1)
- Etudier les effets compensatoires (forcer la forme articulatoire du début de l'inversion)
- Comparer les trajectoires réelles à celles obtenues par inversion
  - dépouillement d'un grand film aux rayons X
  - construction du modèle lié au locuteur du film
  - inversion
    - préférences du locuteur
    - tests de reco.